We would like to thank the reviewer for the valuable comments of our manuscript and constructive suggestions, which help us to improve the quality of the paper. Our responses to specific comments are below. The reviewer's comments are in black and our answers are in blue.

Overall

I would like first to thank the authors for responding to my comments on the initial submission. The manuscript is improved compared to the original submission. However there is still missing information which is important to understanding and interpreting the results.

There are a couple of matters arising from this I'd like to address before moving onto more detailed comments. I now understand that you used different RTTOV coefficients because you had new information about the instrument. However the reader does not know this, nor does the reader know what this new information is, so do not know how relevant these results are to other results. At a minimum you should show the difference between your MWTS-2 RTTOV calculations and those from the coefficients provided by RTTOV itself from the NWPSAF. The value of publishing a paper like this is that it is interesting to see whether different centres get similar or very different impacts from new data. But the implementation described is very cautious and crucial information such as observation errors is missing, so it is not possible to know whether the results (neutral unless other satellite data is removed) were to be expected or not.

Reply:

Currently, only MWTS-2 coefficients for RTTOV v11 are provided on the NWPSAF website. We have downloaded it. However, in our GRAPES system, the RTTOV version is 9.3. The coefficients from NWPSAF is calculated on more vertical levels. The file content is very different from the coefficients file used in RTTOV v9.3. Thus, it cannot be used in RTTOV v9.3. I am sorry we cannot compare the simulation differences between the coefficients from NWPSAF and that from National Satellite Meteorological Center (NSMC) of CMA for they are used in different version of RTTOV.

In fact, since 2014, ECMWF, Met office, Numerical Weather Prediction Center of CMA and NSMC of CMA has set up close collaboration. We have had three teleconference in the last two years. During the two years' cooperation, some calibration problems were found in 2014. After improving the calibration method, the coefficients are calculated again by Prof. Lu Qifeng. Then he provided the coefficients to us to put into the RTTOV v9.3 in GRAPES. So we can have the latest version of coefficient file.

Pro. Lu Qifeng has been to ECMWF twice and evaluated the data quality of MWTS-1 and MWTS-2 using ECMWF NWP model (Lu et al., 2010, 2012, 2015). He calculated the MWTS-2 coefficients for RTTOV v9.3 and RTTOV v7 and renewed it many times. The renewed coefficients are shared in the collaboration group. He also provided the MWTS-2 instrument parameters to ECMWF and Met office to train the coefficients for higher version RTTOV. The coefficients for RTTOV v93 provided by Pro. Lu Qifeng is basically consistent with the coefficient for RTTOV v11 from NWP SAF. Generally, the coefficient on the NWP SAF website maybe not the latest version for MWTS-2, but the simulation capability will be very close to that calculated by Pro. Lu Qifeng, only subtle differences.

Pro. Lu Qifeng have recalculate the coefficient after they improved the calibration method. Details about how to improve the calibration method involves a lot of technical details. It is important but it is not the focus of this article. If the reviewer have more questions about the coefficients. You may contact Pro. Lu Qifeng (luqf@cma.gov.cn), he can give more detailed answers.

References:

Lu, Q., Bell, W., Bauer, P., Bormann, N., and Peubey, C.: An initial evaluation of FY-3A satellite data, ECMWF Technical Memoranda Number 631, ECMWF, Shinfield Park, Reading, UK, 58, 2010.

Lu, Q. and Bell, W.: Evaluation of FY-3B data and an assessment of passband shifts in AMSU-A and MSU during the period 1978-2012, Interim report of Visiting Scientist mission NWP_11_05, Document NWPSAF-EC-VS-023, Version 0.1, 28, 2012.

Lu, Q., Lawrence, H., Bormann, N., English, S., Lean, K., Atkinson, N., Bell, W., and Carminati, F. : An evaluation of FY-3C satellite data quality at ECMWF and the Met Office, ECMWF Technical Memoranda Number 767, ECMWF, Shinfield Park, Reading, UK, 2015.

Thank you for reminder. We forget to mention some important information. Observation errors have been added. The observation error was estimated using statistics of observed-minus-forecast radiance departures. Currently, fixed values of 0.3 K, 0.35 K, 0.35 K and 0.35 K are applied for MWTS channels 5-8 respectively. We have added it in the text. Please see the text (Page 13 Lines 12-14).Quality control scheme of AMSU-A are also added, Please find it in the text (Page 14 Line 29-Page 15 Line 19). More details will be described in the following replies.

Details

Page 3. MWTS-2 does not have better spectral resolution than MWTS-1. It has more channels. The bandwidth of the channels is similar. Therefore the spectral resolution is the same. Its measuring more spectrum. Also note you repeat exactly the same comment on line 2 and 4. The whole of this paragraph is too verbose and repetitive. Its sufficient to say it has more channels, as shown in Table 1, and samples 90 times rather than 30 in each scan line. It is worth discussing the impact of this on integration time and noise. The high sampling, high noise, would produce similar noise characteristics to a low sampling, low noise system like AMSU-A.

Reply:

Thank you for reminder. Originally, we mean that MWTS-2 measures more spectrum than MWTS-1 here, but the "the spectral resolution" is not used properly. We have deleted "The spectral resolution of MWTS-2 is much higher than that of MWTS-1."

We have modified the paragraph to make it more concise and clear. Only one sentence is remained "It is anticipated that the MWTS-2 data could also be useful for NWP modeling systems". Please see the text (Page 3 Lines 17-18).

Page 3 Lines 22-23. The claim that MWTS-2 is "much better than AMSU-A" is not correct. It has one more channel (note statement MWTS-2 has more channels on line 7 can also mislead – it is true, it has one more, but the value of the 51 GHz channel remains unproven, and you certainly don't use it!).

Reply:

Thank you for reminder. We want to say that MWTS-2 is designed to have one more channel and samples more times than AMSU-A. But indeed, it does not mean that it can be better than AMSU-A. The description has been deleted. The statement "MWTS-2 has more channels" is also modified.

In addition, the shortcoming of MWTS-2 compared with ATMS and AMSU-A are also added. "However, it should also be noticed that compared with ATMS, MWTS-2 contains only the 50 GHz frequencies and thus the other frequencies cannot be used for quality control. In addition. Compared with AMSU-A, MWTS-2 does not have the three channels, 23.8, 31.4 GHz and 89 GHz which are often used for cloud detection (Weng and Grody, 1994; Grody et al., 2001; Klaes and Schraidt, 1999). It will make us difficult to detection cloud and precipitation". Please see the text (Page 3. Lines 11-16).

Thank you for pointing out the valuable research direction. We will try to do some detailed studies about the 51 GHz observations later.

Page 5 line 29-31 As stated the use of your own coefficients is OK so long as you are absolutely clear what it is that you have assumed that is different to the official RTTOV coefficients. This has to be clear so the reader can appreciate exactly what is the difference between your modelling of MWTS-2 and other centres modelling of MWTS-2. How different are your results to the official RTTOV coefficients?

Reply:

Please see the reply to the first comment.

Page 6: Your channel selection is very conservative. I am not aware of centres that do not use AMSU channel 5 over the ocean, which is equivalent to MWTS-2 channel 4.

Reply:

Thank you for your suggestion. Yes, we are very conservative in the choice of channels. As the first step, we want to use the data in our NWP model and we hope they do not make our model worse. We will do some research on it and try to use channels 5 in our next study.

Page 6: Why are the plots using CRTM when your assimilation is using your own coefficients, implemented in RTTOV? So the plots do not reflect what you are actually doing? How different are the CRTM simulations to your simulations?

Reply:

I am sorry that the description in this part is not clear. Section 4.2.1 introduce an evaluation of MWTS-2 data. It is just an evaluation before the data assimilation. GFS field and CRTM simulations were not used in the assimilation of MWTS-2 in the global GRAPES system. Section 4.2.2 extract the striping noise from the observations from MWTS-2, then the data are used in the data assimilation.

In section 4.2.1, global simulations of brightness temperature are used as a “reference” for examining the performance of the MWTS-2 instrument. The 6-hour forecasts of the vertical profiles of temperature, specific humidity and the surface pressure from the NCEP global forecast system (GFS) are used as input to CRTM. We also compare the brightness between observations and simulations (O-B) with those of ATMS. ATMS is used here to compare with the MWTS-2 observations. The NCEP GFS background fields and CRTM have been used by Guan et al., (2011), Zou et al. (2011) and Qin et al. (2013) to evaluate the data quality of MWTS-1, MWHS-1 (onboard FY-3A/B) and ATMS (onboard SNPP).

In this section, the initially assessment shows that there are striping noises in MWTS-2 data. Then the striping noises are extracted from the data using the method proposed by Qin et al. (2013). 2013, Qin et al. evaluate the quality of the brightness temperature measurements from ATMS and the NCEP GFS forecast fields and CRTM are used in the simulations. To make our evaluation of MWTS-2 comparable to the results from Qin el al. (2013), we also choose GFS forecasts fields and CRTM in the simulation of satellite radiance.

To make it more clearly, we changed the title of 4.2.1 to “Evaluation of MWTS-2 data quality”. We also added a statement “The global observed brightness temperatures of MWTS-2 channels 5-8 are assessed before they are assimilated in GRAPES”. Please see the text (Page 6. Lines 18-19).

We compare the brightness temperature simulated by different radiative transfer models. The CTRM simulations and the RTTOV v9.3 (using the coefficients provided by Pro. Lu Qifeng (NSMC)) are shown in figures 1-2. Both the radiative transfer model use the GFS NCEP model. There are some subtle differences. In general, the simulation are similar.

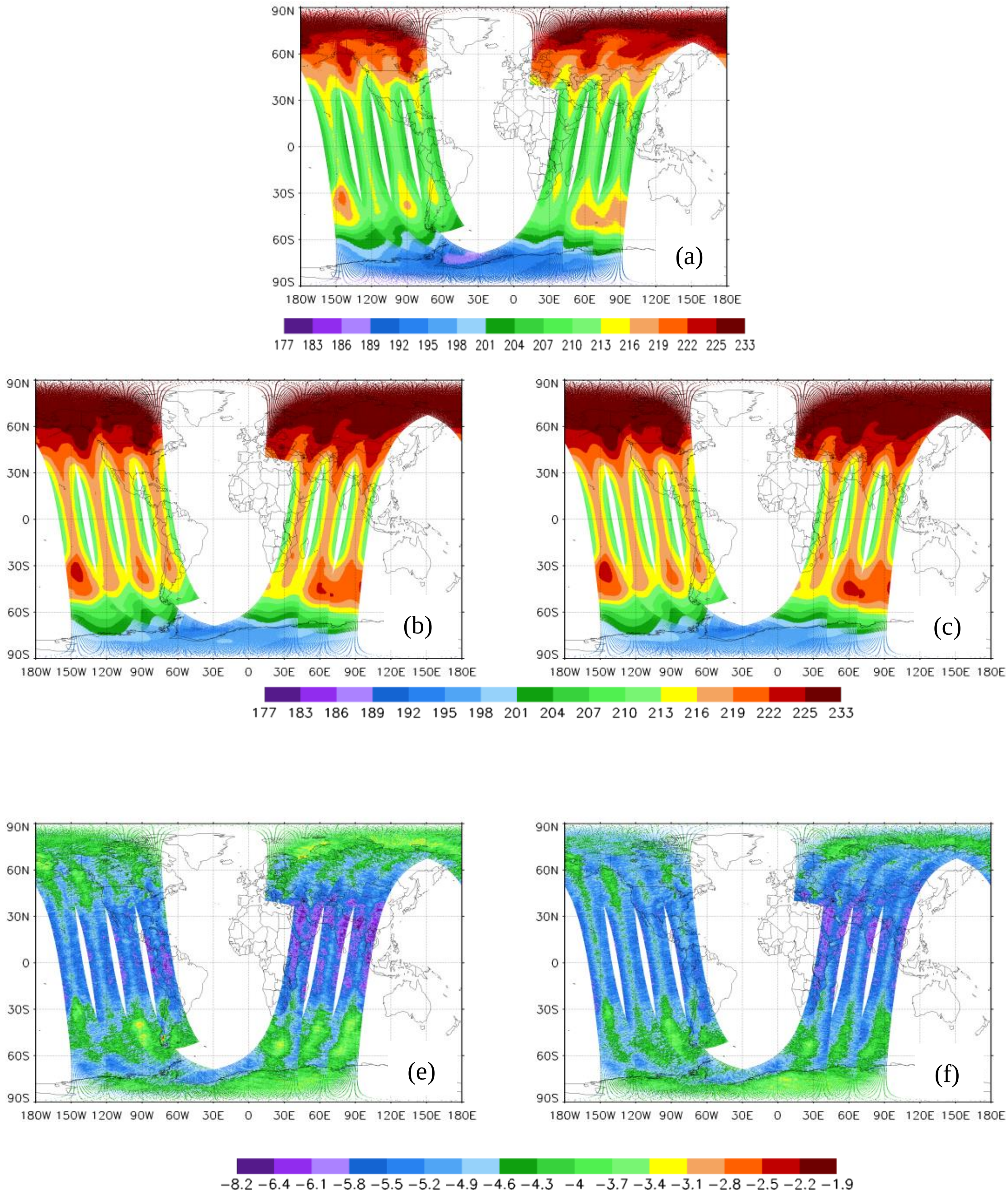

Figure 1. The (a) observations, model simulations using (b) CRTM and (c) RTTOV v93, O-B using (e) CRTM and (f) RTTOV v93 of MWTS-2 channel 7 during 0300-0900 UTC 1 July 2014.

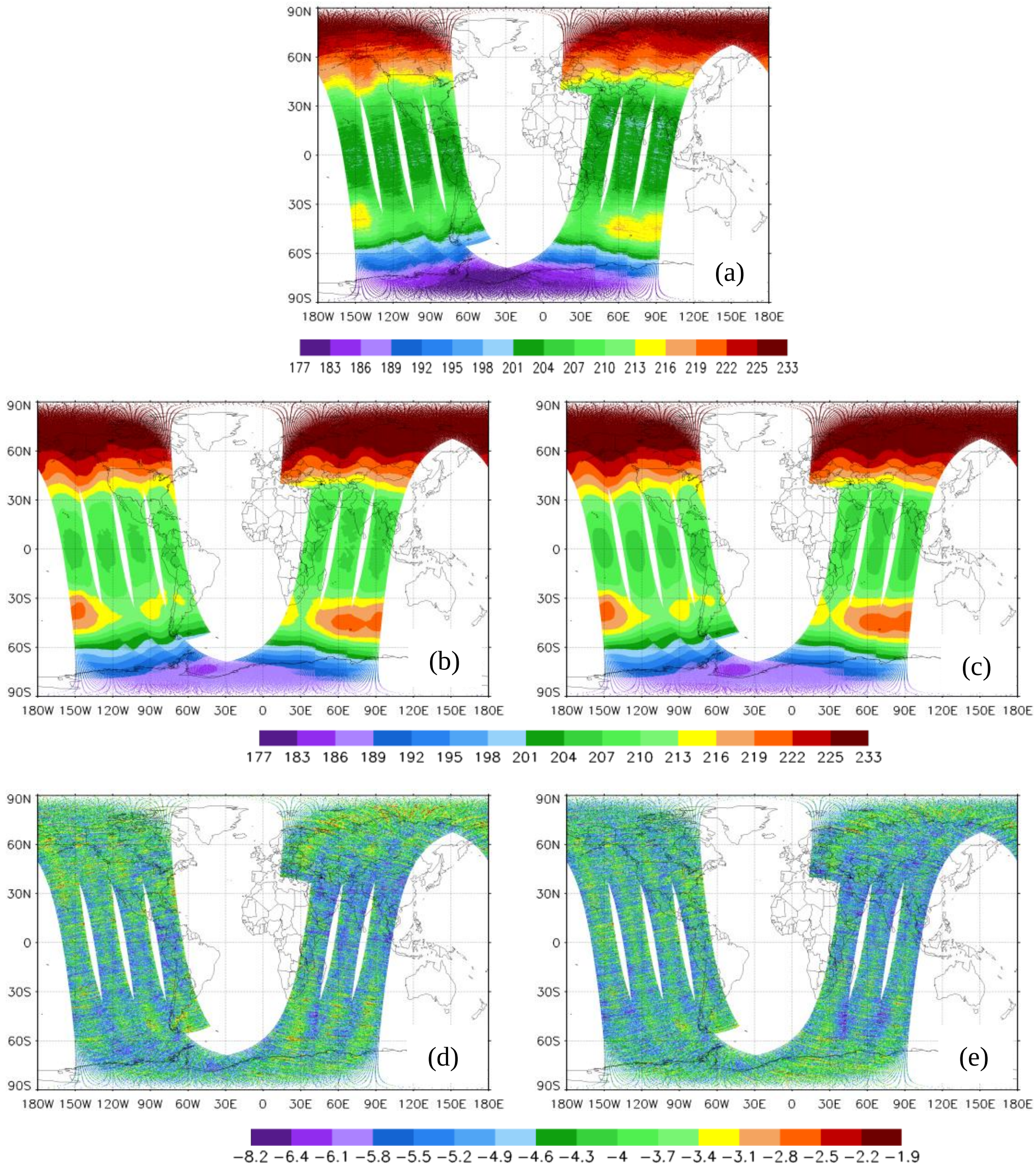

Figure 2. The (a) observations, model simulation using (b) CRTM and (c) RTTOV v93, O-B using (e) CRTM and (f) RTTOV v93 of MWTS-2 channel 8 during 0300-0900 UTC 1 July 2014.

References:

Guan, L., Zou, X., Weng, F., and Li, G.: Assessments of FY-3A Microwave Humidity Sounder measurements using NOAA-18 Microwave Humidity Sounder, J. Geophys. Res., 116, D10106, doi:10.1029/2010JD015412, 2011.

Qin, Z., Zou, X., and Weng, F.: Analysis of ATMS striping noise from its Earth scene observations, J. Geophys. Res. Atmos., 118, 13214-13229, doi:10.1002/2013JD020399, 2013.

Zou, X., Wang, X., Weng, F., and Guan, L.: Assessments of Chinese FengYun

Microwave Temperature Sounder (MWTS) measurements for weather and climate applications, J. Atmos. Ocean. Technol., 28, 1206-1227, 2011.

Page 9: I got very confused reading this. I do not understand Figure 6. Figure 6 tells me it shows me where NOAA-18 AMSU-A LWP product says there is LWP < 0.3 kg/m2 and where the VIRR says its less than 76% cloudy. I assume the white is where NOAA-18 AMSU-A LWP > 0.3 kg/m2 and VIRR says its more than 76% cloudy. I find this figure very difficult to interpret as to how reliable your cloud screening is - but I will do my best to interoret. I understand the figure to show that your method rejects too much data (too few blue points). Its clear VIRR will have many fovs where there is high level ice cloud that is not radiatively important for MWTS-2. So its not a surprise that it over rejects. We also know there will be some occasions when there is low level high liquid water cloud that is radiatively important for MWTS-2 but will not be seen by VIRR, though this will be far less common. So I do not understand the motivation for using VIRR for the cloud screening versus using the lower peaking MWTS-2 channels that you do not assimilate because they are too sensitive to the factors (cloud, surface) that you are trying to QC in the higher peaking channels. This sensitivity makes them very useful for QC. If you think there are specific situations where this is inadequate (perhaps based on comparing this approaching with other approaches for ATMS and AMSU-A) then please document this in the paper. At least tell us why you chose what appears to be a surprising and inappropriate choice of QC for a microwave sounder.

Reply:

For microwave temperature satellite measurements, the two weak water absorption channels (23.8 And 31.4 GHz) and the channel sounding of the scattering process (89 GHz) are often used for cloud detection (Weng and Grody, 1994; Grody et al., 2001; Klaes and Schraidt, 1999). However, the first generation microwave sounding unit on board FY-3A/B has not these channels. It has only has 4 channels. This situation makes it difficult to detect clouds and precipitation by itself. To solve this difficulty, in 2013, a VIRR cloud detection method is proposed for FY-3A/B MWTS-1 (Li and Zou, 2013; Li and Liu, 2015). It can detect cloud contaminated FOVs efficiently (However, it will reject too much data currently. The threshold can be studied and adjusted).

In Sept, 2013, FY-3C was launched, but the second generation microwave sounding unit (MWTS-2) also have not these channels. So we still use the VIRR method in this research. In Figure 6, The CLW method is used here as a reference (Weng and Grody, 1994; Ferraro et al., 2005). We compare the clear region detected by VIRR method and CLW method. We want to show in this figure that the clear FOVs of VIRR method are located in the clear region detected by CLW method though VIRR method rejects too much pixels. Thus, clear pixels detected by the VIRR method is reliable.

Currently, we have not find a good way to use lower peaking MWTS-2 channels to detect the cloud. We have tried to use the O-B method and set a threshold to detect the cloud. However, the surface simulations are not good in GRAPES. In some regions, it will detect a lot of wrong cloud. So we do not use the lower peaking channels and continue to use the cloud detection used by MWTS-1 onboard FY-3A/B.

Last year, when we visit NCEP, Pro.Weng Fuzhong also give us some advisement. He advise us to use the channels 1 and 4 (50.3 GHz and 53.596 GHz) to detect the cloud (Weng and Zou, 2014). We are trying this method now.

Reference:

Weng, F. and X. Zou:30-Year atmospheric temperature record derived by one dimensional variational data assimilation of MSU/AMSU-A observations. Climate Dynamics, 43(7-8):1857-1870, 2014.

Page 9 How do these QC approaches compare to your use of ATMS and AMSU-A. Are you basing these on the window channels? What is the difference in the amount of data passing QC? In general it would be useful to have clear information on the differences between how AMSU-A and ATMS (temperature sounding channels) are assimilated compared to MWTS-2 – is it equally cautious? Are observation errors similar? How does the impact measured here compare to known impact of AMSU-A and ATMS?

Reply:

Thank you for your comments. I would first explain that ATMS is not assimilated here. It is only used to compare with the MWTS-2 observations in Section 4.2. We have added some statements in the text (Page 8 Lines 5-6) to make it clear. The Types of observational data assimilated in GRAPES are shown in table 3.

The quality control scheme of AMSU-A are added in the text. There are also some procedures based on the window channels. Scattering index method is used in the cloud detection of AMSU-A. The observations error are also listed in Table 4. about 27%, 28%, 37%, 73%, 73%, 73% and 70% of the observations are maintained for channels 5-11 after these QC procedures. More channels and more data are used for AMSU-A. Considering there are three AMSU-A instruments used (NOAA-15, NOAA-18 and MetOp-A), the total amount of AMSU-A observations assimilated is about 6.7 times of that of MWTS-2. Please find more details in the text (Page 14 Line 29-Page 15 Line 19, Page 15 Lines 25-27).

Page 10 I think you need to justify your other QC decisions because these are stricter than most centres use. On what basis did you decide not to use channel 5 over sea ice? Did you have evidence of higher departures? Similarly why 500m? What was the basis for the decisions you took? I am amazed how little use you make of channels 5 and 6, these are not low peaking channels. Also the point at which you decide not to use higher peaking channels has not been explained.

Reply:

Thank you for your attention to the details. When we are going to use a new satellite in GRAPES, we are always cautions at first. After the data can be assimilated safely (the assimilation of the data will not ruin the NWP system), we will try to use more data. The quality control scheme of AMSU-A are added in the text. Some of the QC procedures are similar to MWTS-2 but in all, the QC scheme of AMSU-A are not so strict as MWTS-2. More channel 6 (similar to MWTS-2 channel 5 in frequencies) observations are used. AMSU-A channels 5 (53.596 GHZ) observations are also used from AMSU-A. Please see in the text (Page 14 Line 29-Page 15 Line 19, Page 15 Lines 25-27).

In this initial assimilation of MWTS-2 data, channel 5 are only used over ocean. The first reason is that the O-B departures of channel 5 over sea ice (over Antarctic and artic) and land is larger than those over the ocean. The second is that the weighting function (Figure 1) of channel 5 shows that there are some surface signals observed by channel 5. However, simulations of surface signals are not good for GRAPES. Surface channels of AMSU-A (all data of channels 1-4, channel 5 data over land and sea ice) are still not used. Thus, we decided to use only the safest data (over ocean). We do need to do more careful research on the channels selection. This year, we are going to do some research on the simulations of surface channels and do more in-depth study.

We do not used the channel 6 data with terrain higher than 500 m. The reason is similar to those described in the last paragraph. Larger O-B difference are found on the regions with high terrain (eg: Tibetan Plateau). The weighting function also shows that surface signals of high terrain will be probed by channel 6. So we decided not to use these data. The number of 500 is based on the previous experience. The quality control of channel 7 observations of AMSU-A and channel 3 of MWTS-1 onboard FY-3A/B adopts this threshold, so we still use it to maintain the consistency. There are some references about the quality control of AMSU-A and MWTS-1 assimilation in GRAPES (Zhang, 2004; Liu, 2007; Xue, 2008; Zhu, 2008, Li, 2012).

We do not use the higher peaking channels 9-10. In some area, the model top can be as low as 7-8 hPa, there are some signal from the higher atmosphere. In the initial experiments, we decided not to use channel 9 and 10. We are really too conservative in the choice of channels. We will try to use channels 9 and 10 in our next study.

We have added explanations in the text. Please see the text (Page 10 Lines 9-18, Page 6 Lines 9-14).

References:

Zhang Hua, Xue Jishan, Zhu Guofu, et al., 2004: Application of direct assimilation of ATOVS microwave radiances to typhoon track prediction. Adv. Atmos. Sci., 21(2), 283-290.

Liu ZhiQuan, Zhang F Y, Wu X B, Xue J S (2007). A regional ATOVS radiance-bias correction scheme for radiance assimilation. Acta Meteorologica Sinica, 65 (1): 113-123 (in Chinese)

Xue JiShan, Zhuang S Y, Zhu G F, Zhang H, Liu Z Q, Liu Y, Zhuang Z R (2008). Scientific design and preliminary results of three-dimensional variational data assimilation system of GRAPES. Chinese Science Bulletin, 53(22): 3446-3457

Zhu Guofu, Xue Jishan, Zhang Hua, et al., 2008: Direct assimilation of satellite radiance data in GRAPES variational assimilation system. Chinese Science Bulletin, 53(22), 3465-3469.

Li Juan. and Zou, X.: A quality control procedure for FY-3A MWTS measurements with emphasis on cloud detection using VIRR cloud fraction, J. Atmos. Ocean. Technol., 30(8), 1704-1715, doi:10.1175/JTECH-D-12-00164.1, 2013.

Analysis section: This was mostly OK. I would like significance testing on the verification of analysis vs NCEP in Figs 12+13, as is done in Fig. 14. I was not clear in the verification in Fig. 14 was also against NCEP or whether each experiments is verified against its own analysis. This is important to interpretation and must be stated clearly. I would find Fig. 12+13 if it was plotted as a percentage chain in RMS from the control, rather than absolute figures. A 1% change would be important, but a 1% change would not be visible on these plots. Also I found it hard to see a difference between the thin blue and thin black line.

Reply:

Thank you for your suggestion. We have added significance test on the verification against those of the NCEP analysis in Figures 12-13 and Figures 15-16 as suggested. The color of the lines have also been changed.

Figures 18-19 (They are Figures 14-15 in the original manuscript) were verification against their own analysis. The descriptions about the verification method and the figures are in Page14 Line 10.

We also added Figures 14 and 17. They shows the reduction percentages of the impact experiment in RMS from the control.

Please see in the text.

Nice result with the striping noise correction showing positive impact. Personally I'd not call your process “extracting noise”. This sounds strange. I'd say “correction for the striping noise”.

Reply:

We have changed all the “extracting noise” to “correction for the striping noise”. Please see the text.

Conclusions: The main conclusion is that MWTS-2 is neutral when you have AMSUA and ATMS. Given the very cautious use of the data I do not find this surprising, though I don't have all the information I need to know what I expect the impact to be. The positive impact when other satellite data is not used shows that the assimilation system is basically sound, and this is important for CMA, but not so interesting to the community in general. I really felt the paper lacked a proper assessment of what we would expect the impact to be and whether what we measured fell in this range.

Reply:

Thank you for your comments. I would first explain that ATMS is not assimilated here. It is only used to compare with the MWTS-2 observations in Section 4.2. We have added some statements in the text (Page 8 Lines5-6) to make it clear.

The MWTS-2 impact is neutral when AMSU-A is used. When we are going to use a new satellite in GRAPES, we are always cautions at first. When we assimilate MetOp-A AMSU-A data, we are also cautions initially. After a certain period of testing, we begin to use more channels and more data. The MWTS-2 observations is new for us. We need to evaluate the data and become familiar with it gradually. When we obtain the MWTS-2 observations, there are some problems found to be existed in the MWTS-2 data, such as inaccurate land sea masks and stripping noise. The land sea mask problems are solved after NSMC of CMA adjusted the geolocation using a higher resolution map (it is not mentioned in the manuscript). Obvious stripping noise are found to be existed in the MWTS-2 observations (The noise is corrected using PCA+EEMD method in this manuscript). MWTS-2 observations have different characteristic with AMSU-A data. So we are conservative in the quality control scheme in the initial experiments. As the first step, we hope the new observations will not ruin the NWP system (GRAPES). The neutral results of MWTS-2 when AMSU-A are used are in line with our expectations. After all, when a lot of similar observations (AMSU-A) have been assimilated, the impact of new data will not be so significant. In addition, less channels are used for MWTS-2 than AMSU-A. Thus, the impact results are not surprising. According to the reviewer's suggestion, we will try to use more channels and improve the quality control scheme this year.

Overall I think the authors need to provide: - Complete information on what they have assumed about MWTS-2, including differences from assumptions made e.g. by RTTOV - Clear justification for the QC decisions taken - Observation errors – Significance testing on all results - Some indication of what impact would be expected, given impact of other similar data - Some indication therefore whether the impact is as expected, smaller or larger.

Reply:

We have added more descriptions about the MWTS-2 quality control scheme (Page 10 Lines 9-18, Page 6 Lines 9-14), observations errors are shown in a new table (Table 3). We have also added the significance testing on all test (Figures 12-19). The quality control scheme and impact results of AMSU-A are also provided ((Page 14 Line 29-Page 15 Line 19, Page 15 Lines 25-27). The discussion about the impact of MWTS-2 are in the Summary section (Page 17 Lines 3-11).

Please see the text.